STA 2023   Practice Questions   Final Exam

STATISTICAL INFERENCE TABLE: (GIVEN on the test)

Case	parameter	estimator	standard error	Sampling Distribution
one mean	μ	$\overline{x}$	$s/\sqrt{n}$	t (n-1)
Mean of matched pairs difference	μ_d	$\overline{x}_d$	$s_d/\sqrt{n}$	t (n-1)
difference of two independent means	$\mu_1 - \mu_2$	$\overline{x}_1 - \overline{x}_2$	$\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$	t with conservative df = smallest of (n$_1$-1) and (n$_2$-1)
one proportion	p	$\hat{p}$	CI: $\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ ST: $\sqrt{\dfrac{p_0(1-p_0)}{n}}$	z
difference of two independent proportions	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	CI: $\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ ST: $\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$	z

McNemar $TS = \dfrac{YN - NY}{\sqrt{YN + NY}}$ **OR** $TS = \dfrac{NY - YN}{\sqrt{YN + NY}}$

NEED TO KNOW:

Confidence Interval: estimator +/- (t or z) est. standard error

Test Statistic: $(t \quad or \quad z) = \dfrac{estimator - \#\ from\ H_o}{estimate\ of\ stderr}$

Questions 1 – 4 An education researcher has developed a new technique to teach Spanish to high school students. To prove this new method is better, she will teach two groups of students for an entire semester, one with the new method and one with the standard method used today in high schools. She wishes to obtain 12 sets of <mark>identical twins</mark>, 24 people total, and split the twins up so that one of each twin is in Class 1, which will receive the new technique, and one is in Class 2, which will receive the old technique. At the end of the semester both classes will take a standardized exam and the results will be compared.

1. What method should be used to analyze this data?
a) one mean
b) two independent means
c) matched pairs → identical twins
d) two independent proportions

[handwritten: μD]

2. What should our hypotheses be to test if your new method is better than the old method?
a) $H_0: \mu_1-\mu_2 = 0$ vs. $H_A: \mu_1-\mu_2 < 0$
b) $H_0: \mu_D = 0$ vs. $H_A: \mu_D < 0$
c) $H_0: \mu_1-\mu_2 = 0$ vs. $H_A: \mu_1-\mu_2 > 0$
d) $H_0: \mu_D = 0$ vs. $H_A: \mu_D > 0$

[handwritten: group 1 - group 2   90 - 80 = 10]

3. What distribution would you use to look up the p-value for this test of hypothesis?
a) t(11)
b) t(12)
c) t(23)
d) Z

[handwritten: each group has 12 ppl   n=12   df = 11]

4. Suppose we collect the data from this experiment and we find a p-value of 0.009. Which of the following is our conclusion at any reasonable alpha level?
a) There is not evidence of a difference among the two educational methods.
b) There is evidence that the old method is better than the new method. ✗
c) There is evidence that the new method is better than the old method.
d) We cannot determine without more information.

[handwritten: small → YES difference]

Questions 5-7 use the following modifications to the above scenario.
Unable to obtain the 12 sets of twins to conduct the experiment, the researcher has to use two existing Spanish classes at a local high school, one with 12 students and the other one with 11 students.

5. What method should be used to analyze this data?
a) one mean   b) two independent means $\mu_1 - \mu_2$
c) matched pairs   d) two independent proportions $\hat{p}0 - \hat{p}0$

6. What should our hypotheses be to test if your new method is better than the old method?
a) $H_0: \mu_1-\mu_2 = 0$ vs. $H_A: \mu_1-\mu_2 < 0$ b) $H_0: \mu_D = 0$ vs. $H_A: \mu_D < 0$
c) $H_0: \mu_1-\mu_2 = 0$ vs. $H_A: \mu_1-\mu_2 > 0$ d) $H_0: \mu_D = 0$ vs. $H_A: \mu_D > 0$

7. What distribution would you use to look up the p-value for this test of hypothesis?
a) t(11)   b) t(10) $n_1 = 12$ $n_2 = 11$
c) t(23)   d) Z $df_1 = 11$ $df_2 = 10$

8. Find the t* value for a 99% confidence interval.
   a.) 3.169 $df = 10$
   b.) 2.576
   c.) 3.250
   d.) 2.821

Questions 9-11 We suspect that men are convicted of DUI (driving under the influence of alcohol or drugs) more often than women. Let p_1 be the proportion of males convicted of a DUI offense and p_2, the proportion of females convicted of a DUI offense.

9. If the 95% CI for $p_1 - p_2$ is (-.163 , -.02). Which of the following can we conclude?
a) There is evidence that males have a higher proportion of convictions for DUI than females.
b) There is evidence that females have a higher proportion of convictions for DUI than males.
c) There is not evidence that there is a difference in the proportion of convictions for DUI for males and females.   CI doesn't include 0

10. Suppose we test $H_0: p_1-p_2 = 0$ vs. $H_A: p_1 - p_2 \neq 0$. What can we say about the p-value?
a) it is greater than .05   b) it is less than .05
c) it is less than .025   d) cannot say anything about the p-value

11. Are there any problems with the assumptions necessary for these conclusions to be valid?
a) The data may not have been randomly selected, and thus, not representative of the population. ✓
b) The sample sizes may not have been large enough for the formulas to be appropriate. ✓
c) There may be problems with both randomness and sample sizes.
d) There are no problems with this data.

Questions 12 – 14 Many university courses require students to type their written assignments. Do students get better at typing because of this? Researchers want to see if incoming freshmen have a slower average typing rate (in words per minute) than graduating seniors. A random sample of 50 freshmen and 50 seniors was taken, and each was given a typing test to determine their typing rate. This data was entered into Minitab, and the results are shown below:

· independent
· means

```
            N       Mean      StDev    SE Mean
Freshmen   50       27.9      13.4       1.9
Seniors    50       34.0      12.5       1.8

95% CI for mu Freshmen - mu Seniors: ( -11.2,  -0.9)
T-Test mu Freshmen = mu Seniors (vs <): T = -2.32   P = 0.011   DF = 97
```

doesn't include 0

small
rej. H_0: $\alpha = 0.1, 0.05$

12. Which of the following can we conclude from this output (at the 95% confidence level)?
    a) There is evidence that freshmen have a lower average typing rate than seniors.
    b) There is evidence that freshmen have a higher average typing rate than seniors.
    c) There is evidence that Freshmen and seniors type at exactly the same rate, on average.
    d) There is not enough evidence to show a difference between the mean typing rates of freshmen and seniors.

13. Which of the following statements are true?
    I. With a p-value equal to 0.011, we are confident that the mean typing rate for freshmen is smaller than the mean typing rate for seniors. ✓
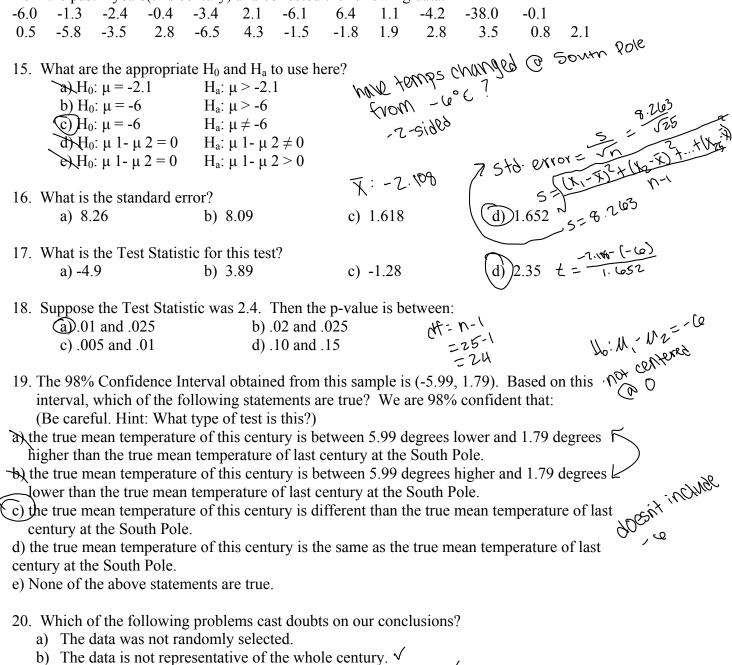   II. The ~~probability~~ *confidence* that the true mean difference for freshmen-seniors lies between -0.9 and -11.2 is .95.
    III. The probability that the mean typing rate of freshmen equals the mean typing rate of seniors is .011. *if H_0 is true*

    a) Only I is true.
    b) Only III is true.
    c) Only I & II are true.
    d) All three statements are true.

14. Which of these are necessary assumptions when making inferences based on this data?
    I. Each sample is randomly drawn from the population ✓
    II. The two samples are independent. ✓
    III. Each population size is at least 10 times larger than the sample size from that population. ✗   *$n \geq 30$*

    a) Only I is assumed.   b) Only III is assumed.
    c) Only I & II are assumed.   d) All three are assumed.

Questions 15-20 Global warming has gotten a lot of media attention over the past few years, but researchers in Antarctica are keeping track of temperatures at the South Pole to try to see if temperatures really are rising. They know that from 1900-1999(last century) the average temperature at the South Pole was -6 degrees Celsius. They have randomly sampled 25 days from the past 2 years(this century) and collected the following data:
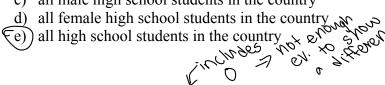
-6.0 -1.3 -2.4 -0.4 -3.4 2.1 -6.1 6.4 1.1 -4.2 -38.0 -0.1
0.5 -5.8 -3.5 2.8 -6.5 4.3 -1.5 -1.8 1.9 2.8 3.5 0.8 2.1

15. What are the appropriate H_0 and H_a to use here?

    a) H_0: $\mu = -2.1$ H_a: $\mu > -2.1$

    b) H_0: $\mu = -6$ H_a: $\mu > -6$

    c) H_0: $\mu = -6$ H_a: $\mu \neq -6$

    d) H_0: $\mu 1 - \mu 2 = 0$ H_a: $\mu 1 - \mu 2 \neq 0$

    e) H_0: $\mu 1 - \mu 2 = 0$ H_a: $\mu 1 - \mu 2 > 0$

[handwritten: have temps changed @ South Pole from $-6\degree C$? -2-sided]

16. What is the standard error?

    a) 8.26   b) 8.09   c) 1.618   d) 1.652

[handwritten: $\bar{X}: -2.108$; Std. error $= \frac{s}{\sqrt{n}} = \frac{8.263}{\sqrt{25}}$; $s = \sqrt{\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \ldots + (x_{25} - \bar{X})^2}{n-1}}$; $s = 8.263$]

17. What is the Test Statistic for this test?

    a) -4.9   b) 3.89   c) -1.28   d) 2.35

[handwritten: $t = \frac{-2.108 - (-6)}{1.652}$]

18. Suppose the Test Statistic was 2.4. Then the p-value is between:

    a) .01 and .025   b) .02 and .025

    c) .005 and .01   d) .10 and .15

[handwritten: $df = n-1 = 25-1 = 24$]

19. The 98% Confidence Interval obtained from this sample is (-5.99, 1.79). Based on this interval, which of the following statements are true? We are 98% confident that: (Be careful. Hint: What type of test is this?)

    a) the true mean temperature of this century is between 5.99 degrees lower and 1.79 degrees higher than the true mean temperature of last century at the South Pole.

    b) the true mean temperature of this century is between 5.99 degrees higher and 1.79 degrees lower than the true mean temperature of last century at the South Pole.

    c) the true mean temperature of this century is different than the true mean temperature of last century at the South Pole.

    d) the true mean temperature of this century is the same as the true mean temperature of last century at the South Pole.

    e) None of the above statements are true.

[handwritten: $H_0: \mu_1 - \mu_2 = -6$ not centered @ 0; doesn't include -6]

20. Which of the following problems cast doubts on our conclusions?

    a) The data was not randomly selected.

    b) The data is not representative of the whole century. ✓

    c) The data does not seem to come from a Normal distribution. ✓

    d) all of the above

    e) only a and b

    f) only b and c

    g) only a and c

Questions 21 – 23 Is left-handedness associated with gender in some way? Researchers were trying to determine if a connection could be made between the dominant hand and gender in high school students. They randomly sampled 1417 high school students across the country and obtained the following data:

	Male Students(1)	Female Students(2)
Left Handed Students	68	97
Right Handed Students	545	707

21. What is the population of interest here?
   a) all left-handed high school students in the country
   b) all right-handed high school students in the country
   c) all male high school students in the country
   d) all female high school students in the country
   e) all high school students in the country

includes 0 → not enough ev. to show a difference

22. The 95% CI was (-0.0430, 0.0242) What can we conclude about the proportions of left handed students among male and female high school students at the 95% confidence level?
a) The true proportion of male students that are left handed is higher than the true proportion of female students that are left handed.
b) The true proportion of male students that are left handed is lower than the true proportion of female students that are left handed.
c) The true proportion of male students that are left handed is equal to the true proportion of female students that are left handed.
d) We cannot conclude that the true proportion of male students that are left handed is different from the true proportion of female students that are left handed.

23. Which of the following assumptions, necessary for the validity of this test, we suspect was violated in this case?
   a) data may not have been randomly selected
   b) males and females are not independent
   c) samples were too small
   d) none seems to be violated

Questions 24-27 Below is data collected from a random sample of 3,647 American professionals from 10 urban cities. We would like to know if male professionals(1) are more likely to have children than female(2) professionals of the same age (33-38).

Proportions ✓

	WOMEN		MEN		TOTALS	
	Without children	With children	Without children	With children	Without children	With children
18-25	265	62	132	164	397	226
26-32	408	216	178	397	583	613
33-38	276	412	86	563	362	978
39-45	89	294	23	275	112	569
46+	23	145	64	455	87	600

24. This would be best analyzed as a:
   a) Matched pairs
   b) One proportion
   c) Two independent means
   (d) Two independent proportions

25. Which data from the table above would we be analyzing in this case?
   a) all women vs. all men
   (b) women aged 33-38 vs. men aged 33-38
   c) women aged 33-38 vs. totals for all people 33-38
   d) women aged 33-38 with children vs. women aged 33-38 without children

26. Are there any problems with this data that would violate the assumptions?
   a) Yes, there is an outlier.   → not possible
   b) Yes, all the counts are too large.
   c) Yes, the data does not come from a Normal distribution.   → don't need b/c proportions
   (d) No problems.

Minitab gave the following output for this analysis.

men - women

```
95% CI for p(1) - p(2):   (0.206951, 0.297170)
Test for p(1) - p(2) = 0 (vs > 0):   Z = 10.95   P-Value = 0.000
```
↙ V. small

27. What conclusion can we make about professional men and women, aged 33-38, in urban cities?
   a) There is not enough evidence to conclude that professional men are more likely than professional women in this age group to have children.
   b) The proportion of professional men who have children is different than the proportion of professional women in this age group.
   (c) Professional men are significantly more likely to have children than professional women in this age group.
   d) Professional women are significantly more likely to have children than professional men in this age group.

[handwritten top: n = 506]

28. An August 1999 ABC news poll of 506 Americans found that 4% of Americans cited "to go further in space" as their top priority for the new millennium, while 2% said "eliminate world hunger." The 95% confidence interval for the difference in proportions was (-.00098, .04098) *[handwritten: includes 0]*
What can we conclude?

a) There is not enough evidence to prove that more Americans support space advances than a solution to world hunger. *[handwritten: circled]*

b) The proportion of people in the population who support space advances and a solution to world hunger is ~~exactly the same~~. *[handwritten: check mark]*

c) We are 95% confident that more Americans support space advances than a solution to world hunger.

d) We are 95% confident that more Americans support a solution to world hunger than space advances.

Questions 29-35 Which case is it? You can use each answer more than once.

a) one mean
b) one proportion
c) matched pairs with means
d) two independent means
e) two independent proportions

[handwritten: A] 29. A manufacturer of nails is interested in determining if their nails are meeting specifications. A sample of 100 nails was taken. Is the average length of the nails longer than 3 inches? *[handwritten: one, means]*

[handwritten: D] 30. A sociologist suspects that members of pairs of heterosexual siblings of the same gender (brother-brother or sister-sister) get married at earlier ages on average than siblings who come from mixed-gender (brother-sister) pairs. *[handwritten: 2 groups, same diff, means]*

[handwritten: C] 31. Does pesticide X affect the yield (in lbs) of any of 10 different varieties of tomatoes ? Two fields (one treated and one untreated) are planted with 2 rows each of 10 varieties of tomatoes, and their yields are compared. *[handwritten: number not proportion, 2 groups = fields, -dependent, each variety in both fields]*

[handwritten: B] 32. A gourmet pet food shop owner wants to estimate how many of his customers would continue to shop with him if he moved across town. He asks a random 50 shoppers if they'd drive across town to shop with him, and he makes a confidence interval to help make the decision to move or not. *[handwritten: yes/no question, 1 group]*

[handwritten: C] 33. The migration of African buffalo herds might be affected by the weight of the transponder used to track them. Last year, the scientists tried their standard transponder on eight herds of buffalo and recorded how far each group traveled. This year they will swap out the heavy transponders with more expensive, lighter ones and see if the same buffalo herds travel farther than last year. *[handwritten: dependent]*

[handwritten: E] 34. A political action group wonders if college graduates are more likely to support increased penalties for repeat offenders than are those without college degrees. *[handwritten: 2 groups: college grads & ppl w/o college degree, proportion]*

[handwritten: A] 35. Is the ice-cream cup filling machine accurate or not? A random sample of 20 quarts of ice cream is taken to see if the average contents is 247 mL of product. *[handwritten: mean, 1 group]*

C 36. Does eating breakfast improve productivity? In a Guess jeans factory, the workers of one sewing crew of 20 are fed a hearty breakfast for a month and then the same sewing crew of 20 is asked to go without eating until lunch time for another month. The teams' productivity is compared for the two months.

number means *dependent*

E 37. Equal rights activists claim that, if admissions into a college's architecture program are fair, half of all students admitted should be female. To test this hypothesis, they look at admissions data from 1982 (before aggressive recruiting of women) and 2001, asking "Has the proportion of women admitted changed significantly in the last decade?"

↳ 2 groups: 1982 & 2001

General Questions

$0 \le p \le 1$

38. When are p-values negative?
   a) when the test statistic is negative.
   b) when the sample statistic is smaller than the hypothesized value of the parameter
   c) when the confidence interval includes only negative values
   d) when we fail to reject the null hypothesis
   e) never

39. When can we decide to fail to reject the null hypothesis by looking at a confidence interval?
   a) when the interval includes only negative values
   b) when the interval includes only positive values
   c) when the interval includes both positive and negative values
   d) when the interval includes the hypothesized value of the parameter
   e) never

Questions 40 - 44 Two UF students are investigating how long it takes shoppers to find parking spaces at the Publix in Butler Plaza on different days of the week. To collect data, they position themselves at the parking lot entrance closest to Publix and, using a stopwatch, they record how long (in minutes) it takes every fifth car to come to a complete stop in a parking space. "Parkwkda" regards data collected on a weekday; "Parksat" on Saturday, and "Parkgame" on a Saturday which was also a UF football game day.

NOTE: Minitab output appears on the following page. Some are relevant to this question, and some are not. Read them carefully.

40. How long, on average, does it take those in the sample to find a parking space on a weekday?
   a)   0.935 minutes
   b)   0.395 minutes
   c)   0.740 minutes
   d)   cannot be determined from the information given

independent

41. Does it take longer to find a parking space on Saturday than on a weekday? This should be analyzed as: *means*
   a) two independent means   b) two independent proportions   c) matched pairs

42. What can we conclude about the time it takes to find a parking space on Saturday compared to a weekday?
   a)   The pvalue for this significance test is .23, which means there is not enough evidence to prove a significant difference between these two times.
   b)   The pvalue for this significance test is .33, which means there is not enough evidence to prove a significant difference between these two times.
   c)   The pvalue for this significance test is .0005, which means that it takes longer on a Saturday to find a parking space.
   d)   The pvalue for this significance test is .002, which means the time to find a space on Saturday is different than on a weekday.

indep.

43. Does it take longer to find a parking space on a Saturday than on a game day? This should be analyzed as: *2 mean*
   a) two independent means   b) two independent proportions   c) matched pairs

44. What can we conclude about the time it takes to find a parking space on a Saturday compared to a game day?
   a)   The pvalue for this significance test is .23, which means there is not enough evidence to prove a significant difference between these two times.
   b)   The pvalue for this significance test is .33, which means there is not enough evidence to prove a significant difference between these two times.
   c)   The pvalue for this significance test is .0005, which means that it takes longer on a game day than on a Saturday to find a parking space.
   d)   The pvalue for this significance test is .002, which means the time to find a space on a game day is different than on a Saturday.

```
Two Sample T-Test and Confidence Interval
Two sample T for parkwkday vs parksat


              N      Mean      StDev    SE Mean
parkwkda    20      0.395     0.307     0.069
parksat     20      0.935     0.583      0.13
```

95% CI for mu parkwkda - mu parksat: (-0.842, -0.24)
T-Test mu parkwkda = mu parksat (vs <): T = -3.66 P = 0.0005
DF = 28

(handwritten: ← is a diff)
(handwritten: sat = bigger/longer)
(handwritten: ↑ v. small)

```
Paired T-Test and Confidence Interval

Paired T for parkwkday - parksat


                  N        Mean      StDev    SE Mean
parkwkda         20       0.395     0.307     0.069
parksat          20       0.935     0.583     0.130
Difference       20      -0.540     0.678     0.152
```

95% CI for mean difference: (-0.857, -0.223)
T-Test of mean difference = 0 (vs not = 0): T-Value = -3.56 P-
Value = 0.002

```
Two Sample T-Test and Confidence Interval
Two sample T for parksat vs parkgameday


              N      Mean      StDev    SE Mean
parksat     20      0.935     0.583      0.13
parkgame    20      0.740     0.421     0.094
```

(handwritten: ↙ include 0)

95% CI for mu parksat - mu parkgame: (-0.13, 0.522)
T-Test mu parksat = mu parkgame (vs not =): T = 1.21 P = 0.23
DF = 34

(handwritten: ↑ rather large)

```
Paired T-Test and Confidence Interval
Paired T for parksat - parkgameday


                  N        Mean      StDev    SE Mean
parksat          20       0.935     0.583     0.130
parkgame         20       0.740     0.421     0.094
Difference       20       0.195     0.879     0.197
```

95% CI for mean difference: (-0.217, 0.607)
T-Test of mean difference = 0 (vs not = 0): T-Value = 0.99 P-
Value = 0.33

45. Which of the following is TRUE?
   a) For the method of two independent means, the size of the two samples can be different. ✓
   b) Matched pairs design uses the standard normal distribution, since the population is required to be normal. ✓
   c) The method of two independent means can be used for categorical response variables, whereas matched pairs design cannot. ✓
   d) All of the above.

indp means & proportions have v. simlar formulas for calculations

46. Which of the following is FALSE?
   a) We generally prefer the method of two independent means over matched pairs design because the latter uses dependent samples.
   b) We generally prefer matched pairs design over the method of two independent means since matched pairs design benefits from the fact that potential sources of bias are controlled. ✓
   c) We generally prefer matched pairs design over the method of two independent means since matched pairs design often results in smaller standard errors. ✓
   d) All of the above.

Questions 47 – 49 According to a study in 2005 by American Research Group, an American adult spends on average an amount of $942 on gifts for the Christmas holiday. We claim that students at UF are likely to spend less than that amount. To test this claim, we randomly surveyed 25 students across campus, the resulted sample mean is 820.24, and the standard deviation is 260.4.

47. Which of the following hypothesis is correct?
   a) H_o: $\mu = 942$, H_a: $\mu < 942$.
   b) H_o: $\mu = 0$, H_a: $\mu \neq 0$.
   c) H_o: $\mu = 942$, H_a: $\mu > 942$.
   d) H_o: $\mu = 942$, H_a: $\mu \neq 942$.

48. Conduct the significant test, what is the value of the test statistic?
   a) -2.338
   b) 2.338
   c) -0.468
   d) 0.468

$$t = \frac{\bar{X} - \mu_z}{s/\sqrt{n}} = \frac{820.24 - 942}{260.4/\sqrt{25}} = \frac{-121.76}{52.08} = -2.338$$

49. What is the associated p-value?
   a) $0.01 < \text{p-value} < 0.025$
   b) $0.02 < \text{p-value} < 0.05$
   c) $0.005 < \text{p-value} < 0.0125$
   d) $\text{p-value} = 0.0096$

$df = 24$

0.025   0.01

50. A public health clinic administers a survey that addresses patient satisfaction with services. The survey uses a summary score on a 70 point scale, with 70 indicating the highest possible satisfaction. The survey instrument has been tested in other environments and has historically had mean $\mu = 50$.
We seek evidence that this population has mean score μ that is less than 50 and take a SRS of $n = 100$. Suppose this sample gives a mean score of 48.4 and a standard deviation $s = 7.5$
What is the p-value for this test?

$H_0: \mu = 50$

$H_a: \mu < 50$

$t = \dfrac{48.4 - 50}{7.5/\sqrt{100}} = \dfrac{-1.6}{0.75} = -2.13$

$df = 99$

   a) Between .05 and .025
   b) Between .1 and .025
   c) .0166
   d) None of the above

51. Lithium carbonate is a drug used to treat bipolar mental disorders. The average dose in well-maintained patients is $\mu = 1.3$ mEq/L. A random sample of 25 patients on lithium demonstrates a mean lithium level $\bar{x}$ of 1.4 mEq/L and a standard deviation $s = 0.3$ mEq/L. Conduct a one sample t test to see if the observed difference is significant. Use a two-sided alternative, as improper dosing would include both under- and over-dosing. What is the p-value for this test?

   a) Between .05 and .01
   b) Between .1 and .2
   c) Between .05 and .1
   d) Between .01 and .025
   e) None of the above

$t = \dfrac{1.4 - 1.3}{0.3/\sqrt{25}} = \dfrac{0.1}{0.06} = 1.667$

$df = 24$

0.1 ← 0.05 → mult by 2 b/c 2-sided

52. A test of $H_0: \mu = 110$ vs. $H_a: \mu \neq 110$ based on $n = 18$ gives $t_{stat} = 2.11$. What is the p-value for this problem?

two-sided

$df = 17$

0.025

   a) .05
   b) .025
   c) Between .05 and .1
   d) Between .01 and .025
   e) None of the above

53. A sample of birth weights (grams) of infants who had died of Sudden Infant Death Syndrome (SIDS) in a large metropolitan area was {2998, 3740, 2031, 2804, 2454, 2780, 2203, 3803, 3948, 2268}. The mean weight of all births in this metropolitan was 3300 grams. What are the degrees of freedom associated with the one-sample t-test of whether the mean birth weight of SIDS cases is significantly different from that of the rest of the population? (Two-sided test.)

   a) 11
   b) 10
   c) 9
   d) None of the above

$df = n - 1$

$n = 10$

54. Suppose a study to determine if the height of male athletes is larger than the average male height believed to be 70 inches was performed on a sample of 100 athletes. The p-value for this test is .003. Which one of theses statements is *true*?   rej @ all usual & levels
0.1, 0.05, 0.01

   a) We reject the null hypothesis at the $\alpha = 0.1$, 0.05, but not at the 0.01 level

   b) There is a 0.3% probability that we observe as low these assuming the null hypothesis is true

   c) 99.7% of the male athletes from the sample are taller than 70 inches

   d) 99.7% of male athletes in general are taller than 70 inches

   e) There is a 0.3% probability that we observe as low these assuming the null hypothesis is false

Questions 55 -56 The drying time of a certain type of paint under specified conditions is normally distributed with mean 75min. Chemists have proposed a new additive designed to reduce the average drying time. It is believed that drying times will remain normally distributed. The hypotheses to be tested are Ho: $\mu=75$ versus Ha: $\mu<75$. Experimental data from a random sample of 25 test specimens provided an estimate of 9 for the standard deviation.

55. What is the test statistic?

$\mu = 75$
$n - 25$
$s = 9$

$t = \dfrac{\bar{X} - \# \, Ho}{s/\sqrt{n}}$

a) $t = \dfrac{\bar{x}+75}{9/\sqrt{25}}$

b) $t = \dfrac{\bar{x}-75}{9/\sqrt{25}}$

$t = \dfrac{\bar{X} - 75}{9/\sqrt{25}}$

c) $t = \dfrac{\bar{x}+9}{75/\sqrt{25}}$

d) $t = \dfrac{9/\sqrt{25}}{75-\bar{x}}$

56. If the test statistic is equal to -2.80, for what values of α (probability of Type I error) would we reject Ho?

df = 24

$0.001 < p < 0.005$

a) 0.10
b) 0.05
c) 0.01
d) All of the above

Questions 57- 59 A random sample of 31 students were asked two questions:

- What is your gender? Male (1) or Female (2).
- How much money in dollars did you spend on textbooks this semester?

The Minitab output looks like:

```
…  T-Test and Confidence Interval

…  T for Books

Sex            N        Mean      StDev     SE Mean

1             17         325        116         28
2             14       263.9       64.4         17

95% CI for mu (1) - mu (2): ( -7,   129)
T-Test mu (1) = mu (2) (Vs not =): T = 1.85   P = 0.076   DF = 25
```

57. What hypotheses are being tested? Let μ denote the mean for book expenditure.

a) Ho: $\mu_1-\mu_2 \neq 0$ versus Ha: $\mu_1-\mu_2=0$
b) Ho: $\mu_1-\mu_2=0$ versus Ha: $\mu_1-\mu_2<0$
c) Ho: $\mu_1-\mu_2=0$ versus Ha: $\mu_1-\mu_2 \neq 0$
d) Ho: $\mu_1-\mu_2 \leq 0$ versus Ha: $\mu_1-\mu_2>0$

do males v females spend diff amounts on textbooks for a semester?

58. What test is being used?

a) 1-sample t-test. *prop*
b) McNemar. →
c) paired t-test.
d) 2-sample t-test.

money = means
boys v girls = indep. groups

59. At what significance level can we reject the Ho? *p-value*

a) At $\alpha=0.01$ but not at $\alpha=0.10$.
b) At $\alpha=0.05$ but not at $\alpha=0.10$.
c) At all $\alpha=0.01$ $\alpha=0.05$ $\alpha=0.10$
d) A $\alpha=0.10$ but not at $\alpha=0.05$.
e) None of the above.

P=0.076
0.076 < 0.1

Questions 60 and 61: A researcher at a local university wants to find out the relationship between person's height right before sleep and right after sleep. She asks 50 people, randomly selected, to participate in the study. These people monitor their heights for three days. Every night, right before the participants go to sleep, they measure their height and then sleep for approximately eight hours. After they wake up, they measure their height again, and take the difference between the measurements:

Difference = Height After Sleep – Height Before Sleep

This routine is performed for three consecutive nights and three differences are obtained. Finally, the differences are averaged to obtain an average difference, D, between heights. The researcher is using the following hypothesis test in her study

Ho: $\mu_D = 0$
Ha: $\mu_D > 0$ ← one sided

60. What is the researcher trying to show using the above hypothesis test?
   a) People, on average, are taller after sleep than before sleep
   b) People, on average, are shorter after sleep than before sleep
   c) There is a difference between people's average height before and after sleep
   d) There is no difference between people's average height before and after sleep

61. If the p-value for the above hypothesis test is 0.21, what conclusion can we make, with α being the usual 0.05? *← large*
   a) There is evidence that people, on average, are taller after sleep than before sleep
   b) There is evidence that people, on average, are shorter after sleep than before sleep
   c) There is no evidence of a difference between people's average height before and after sleep
   d) There is definitely no difference between people's average height before and after sleep

Questions 62 – 63 After finishing her study of people's heights, the researcher finds an article about the relationship between people's weight before and after sleep. The article describes a study done at one of the universities, where an interesting result has been found. Based on the result described in the article, she wants to test whether students at her university weigh less after night's sleep. Each student is weighed before and after sleeping.

62. Based on researcher's goal, what could be the possible null hypothesis of her hypothesis test? *← no diff*
   a) People weigh more on average after sleep
   b) People weigh less on average after sleep *← Ha*
   c) There is a difference between the people's average weight before and after sleep
   d) There is no difference between people's average weight before and after sleep

63. If the p-value for the test conducted by the researcher is 0.003, α equal to 0.05, what can be said about the conclusion of the test? *[handwritten: α small · less than α = 0.05]*
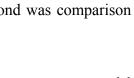   a) There is evidence that students at her university weigh more after sleep than before sleep
   b) There is evidence that students at her university weigh less after sleep than before sleep *[circled]*
   c) There is evidence that students at her university weigh the same before and after sleep
   d) There is evidence that students at her university have different weight before and after sleep
   e) No conclusion can be drawn from the obtained p-value

64. In the problem described for questions 62/63 and for 60/61, what types of analysis were we performing? *[handwritten: dep. groups · same person weighed/measured twice]*
   a) Both tests were comparisons of two independent proportions
   b) Both tests were comparisons of two independent means
   c) Both tests were comparisons of two dependent means *[circled]*
   d) The first test was comparison of independent proportions and the second was comparison of independent means
   e) The first test was comparison of independent means and the second was comparison of dependent means

Questions 65 -70 A poker craze have swept the nation a few years ago, and some experts are concerned that it will lead to an increase in gambling addictions, particularly for adolescents. Each of the five situations presented below describes an inference that we would like to make about a **different** parameter. Match each of the five situations below with the parameter of interest from the list:

[handwritten: hint to use each answer only once]

   a) one mean
   b) one proportion
   c) difference of two independent means
   d) difference of two independent proportions
   e) difference of dependent proportions
   f) mean of matched paired differences

[handwritten: ✓ proportion]

B   65.   A survey finds that 70% of children aged 12 to 17 have gambled in the past year.

A   66.   Although most children start playing just for fun, the average age at which they start gambling for money is around 11 years old. *[handwritten: ↓ mean]*

D   67.   Among 8th graders, 42% of boys and 19% of girls gambled last year.

F   68.   Although some people claim that poker playing helps children with their math skills, a study *[handwritten: mean]*

          found no increase in the average students' math grades from before they started playing. *[handwritten: matched pairs]*

C   69.   Studies have found that boys bet higher amounts of money than girls, on average. *[handwritten: Means]*

E   70.   Even though 79% of children agree that gambling can become addictive, the majority think it could not happen to them. *[handwritten: · same ppl asked both questions — dependent]*
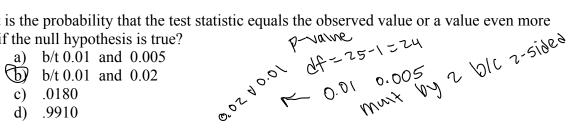
Questions 71 – 75 In some mining operations, a byproduct of the processing is mildly radioactive. Of prime concern is the possibility that release of these byproducts into the environment may contaminate the freshwater supply. There are strict regulations for the maximum allowable radioactivity in supplies of drinking water, namely an average of 5 picocuries per liter (pCi/L) or less. However, it is well known that even safe water has occasional hot spots that eventually get diluted, so samples of water are assumed safe unless there is evidence to the contrary. A random sample of 25 specimens of water from a city's water supply gave a mean of 5.39 pCi/L and a standard deviation of 0.767 pCi/L.

[handwritten: $\mu = 5$, $n = 25$, $\bar{X} = 5.39$, $S = 0.767$]

71. What is the Hypothesis?
    a) $H_0: \mu = 5.39$, $H_a: \mu \neq 5$
    b) $H_0: \mu = 5$, $H_a: \mu \neq 5$ *(circled)*
    c) $H_0: \mu = 5.39$, $H_a: \mu \neq 5.39$
    d) $H_0: \mu = 5.39$, $H_a: \mu \neq 5$
    e) $H_0: \mu = 5$, $H_a: \mu \neq 5.39$

72. What is the test statistic?
    a) $z = 2.542$
    b) $t = 0.508$
    c) $t = 2.542$ *(circled)*
    d) $z = 0.508$

[handwritten: $t = \dfrac{\bar{X} - \# H_0}{S/\sqrt{n}} = \dfrac{5.39-5}{0.767/\sqrt{25}}$; $t = \dfrac{0.39}{0.1534} = 2.542$]

73. What is the probability that the test statistic equals the observed value or a value even more extreme if the null hypothesis is true?
    a) b/t 0.01 and 0.005
    b) b/t 0.01 and 0.02   *(circled)*
    c) .0180
    d) .9910

[handwritten: p-value; $df = 25-1 = 24$; $0.02 < 0.01$; 0.01 0.005; must by 2 b/c 2-sided]

74. What is the 95% confidence interval for the population mean amount of radioactive byproducts?
    a) (5.07, 5.71)   *(circled)*
    b) (5.00, 5.77)
    c) (5.09, 5.69)
    d) (4.92, 5.64)

[handwritten: $\bar{X} \pm t \cdot S/\sqrt{n}$; $5.39 \pm 2.064\left(\dfrac{0.767}{\sqrt{25}}\right)$; 5.39 ± 0.3166 ; $(5.07, 5.71)$; $t: df = 24$ @ 95% Conf ; 2.064]

75. What is the conclusion if you use significance level $\alpha = 0.05$?
    a) We fail to reject that the average radioactivity in the water is 5 pCi/L
    b) We fail to reject that the average radioactivity in the water is 5.39 pCi/L
    c) We have enough evidence to reject that the average radioactivity in the water is 5 pCi/L   *(circled)*
    d) We have enough evidence to reject that the average radioactivity in the water is 5.39 pCi/L

[handwritten: CI doesn't include 5 ; sig. test in #73 ; p-val: $0.01 \leq p \leq 0.02$]

76. What is the hypothesis for the McNemar test?

a) Ho: p1-p2=0, Ha: p1- p2>0.
b) Ho: p1-p2=0, Ha: p1-p2<0.
c) Ho: p1-p2=0, Ha: p1-p2≠0.
d) Ho: p=0, Ha: p≠0.
e) Ho: p=0, Ha: p>0.

77. What is the test statistics for the McNemar test for the table below?

	yes	no
yes	a	b
no	c	d

$$\frac{YN-NY}{\sqrt{YN+NY}} \qquad or \qquad \frac{NY-YN}{\sqrt{YN+NY}}$$

· will give same #
  -one pos. one neg.

$$YN = c$$
$$NY = b$$

a) Test Statistic= $\dfrac{b+c}{b-c}$

b) Test Statistic= $\dfrac{a-d}{\sqrt{a+d}}$

c) Test Statistic= $\dfrac{b-c}{\sqrt{b+c}}$

d) Test Statistic= $\dfrac{a-b}{\sqrt{a+b}}$

e) Test Statistic= $\dfrac{a-d}{\sqrt{a+d}}$

Question 78 – 80 For a random sample of 1600 voting-age Americans, 944 people indicate approval of the President's performance in office. A month later, 880 of these same 1600 people indicate approval. Table below shows the result.

DEPENDENT

First Survey	Second Survey		
	Approve	Disapprove	Total
Approve	794	150	944
Disapprove	86	570	656
Total	880	720	1600

78. Determine whether the sample is dependent or independent?
a) Dependent.
b) Independent.

79. What is the value of the test statistic?
a) Test statistic = 5.25 or -5.25.
b) Test statistic =6.07 or -6.07.
c) Test statistic=4.17 or -4.17.
d) Test statistic=0.271 or – 0.271.

$$YN = 86 \qquad NY = 150$$

$$TS = \frac{150 - 86}{\sqrt{150+86}} = \frac{64}{15.36} = 4.17$$

$$OR \quad TS = \frac{86-150}{\sqrt{150+86}} = \frac{-64}{15.36} = -4.17$$

$$p\text{-val} \sim 0$$

80. Give the conclusion for the test (take α =0.05)
a) We reject the null and conclude that there is no significant evidence of the change in approval rating.
b) We fail to reject the null and conclude that there is significant evidence of the change in the approval rating.
c) We accept the null and conclude that there is significant evidence of the change in the approval rating.
d) We reject the null and conclude that there is significant evidence of the change in approval rating.

NEVER accept Ho

81. In hypothesis testing, if you made a Type I error, you would do which of the following?
a.) Reject Ho when Ho was True
b.) Reject Ho when Ho was False
c.) Fail to Reject Ho when Ho was True
d.) Fail to Reject Ho when Ho was False

H_o : innocent

H_a : guilty

convicted as guilty when actually innocent
· rej. Ho when it's true

82. In hypothesis testing, if you made a Type II error, you would do which of the following?
a.) Reject Ho when Ho was True
b.) Reject Ho when Ho was False
c.) Fail to Reject Ho when Ho was True
d.) Fail to Reject Ho when Ho was False

setting someone free [saying they are innocent] when they are actually guilty
· fail to rej Ho when Ho is false

1. C. In a *matched pairs* study, subjects are matched in pairs and the outcomes are compared within each matched pair. Key word: twins.

2. D. To analyze the data, we first subtract the exam score in Class 1 from that in Class 2 to obtain the difference for each pair of twin. Here μ_D is the mean difference. The null hypothesis says that no difference occurs, and H_a says that exam scores are higher in Class 1. When you subtract class1- class2, if class 1 scores are higher, $\mu_D > 0$.

3. A. In a matched pairs analysis, we assume the population of difference has a normal distribution, so the t procedures can be applied here. And the t distribution is with n-1 (12-1=11) degrees of freedom.

4. C. Use the rule p-value <alpha, reject H_o. Our reasonable alpha levels are .10,.05, and .01. We reject H_0 at all these levels. So, we can conclude the new method is better than the old method.

5. B. The goal of inference is to compare the responses (score mean) in two groups and the responses(score mean) in each group are independent of those in the other group.

6. C. Because H_a expresses the effect we hope to find evidence *for,* so the H_a here should be:$\mu_1 - \mu_2 > 0$. (the mean exam scores in Class 1 is higher than that in Class 2). If μ_1 is greater than μ_2, we could get a value greater than zero. And then set up $H_0 : \mu_1 - \mu_2 = 0$ as the statement that the hoped-for effect is not present.

7. B. By hand, we use the conservative degrees of freedom, the smaller of $n_1 - 1$ and $n_2 - 1$. Here, $n_1 = 12$, $n_2 = 11$, so we choose DF=$n_2 - 1 = 11 - 1 = 10$.

8. A. It follows the t (n-1=10) distribution. For 99% confidence interval, Table D gives t*= 3.169.

9. B. From the CI, we are 95% confidence that the difference in the proportion is between -.163 and -.02. So we can say that females have a higher proportion of convictions for DUI than males. Since p1-p2 is negative, the proportion of group 2 (females) must be bigger than the proportion of group 1.

10. B. From the confidence interval, we know the p1-p2 should be less than 0, which means we should reject the null. Besides, we always reject the null when the p-value is smaller than the alpha. So, the p-value is less than .05 here.

11. C. When we want to compare two proportions, we should have assumptions that

the sample should be randomly selected and the sample size should be large enough.

12. A. From the CI, we are 95% confidence that the difference in the average typing
rate ($\mu_{freshmen}$ - $\mu_{seniors}$) is between -11.2 and $-.9$. So, we can conclude that
freshmen have a lower average typing rate than seniors. Since the results are negative, the
2nd group must be higher seniors. Seniors are faster.

13. A. The p-value is small, so reject H_0. So I is correct. II is not true because if
the endpoints of the CI are given, use the term confidence, not probability. III is also
wrong because the definition of the p-value is the probability that you would see a
result this extreme if the null were true.

14. C. This is a two independent mean problem. A two independent sample mean
problem can arise from a randomized comparative experiment that randomly divides the
subjects into two groups and exposes each group to a different treatment. The
assumptions to those kinds of problems are: each sample is randomly drawn from the
population and the responses in each group are independent of those in the other group.
So I and II are correct.

15. B. Because H_a expresses the effect we hope to find evidence *for,* so the H_a here
should be: H_a : μ >-6. And then set up H_0 : μ =-6 as the statement that the hoped-for
effect is not present. Notice we only have this century's results, so this is a one sample
test.

16. D. Use your calculator to get the standard deviation s first. Then, use the
formula $SE_{\bar{x}} = \dfrac{s}{\sqrt{n}}$ and we can get the answer should be 1.652.

17. D. $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{-2.11-(-6)}{1.652} = 2.35$

18. A. The P-value is the probability of observing a value of t at least as small as the
one that we observed, t=2.35. Based on alternative, we want to shade
greater than 2.35. Table D shows that 2.35 lies between 2.172 and 2.492
critical values of the t (24). So, the p-value is between 0.01 and 0.025.

19. C. Since the CI doesn't include -6, we can only say the true mean temperature of
This century is different than the true mean temperature (-6) of last century at
the South Pole.

20. F. The description states that the data was randomly selected, so A is not a
problem. This century data is only two years. We do not have data for the
full one hundred years, so B is a problem. If you make a dot plot of the
data, you will notice that –38 is an extreme outliers, so there is a problem

with C. So, B and C are problems.

21. E. Since researchers were trying to determine if a connection could be made
     between the dominant hand and gender in "high school students", the
     population of interest should be all high school students in the country.

22. D. Since the CI includes 0, we cannot determine if there's a difference between
    those two proportions. If the difference is < 0, it represents the true proportion of male
     students that are left handed is lower than the true proportion of female students that are
     left handed. If the difference is =0, it means those two proportions are equal.
     Otherwise, it represents the true proportion of male students that are left handed is
     higher than the true proportion of female students that are left handed

23. E. The researchers randomly sampled some high school students across the
     country, so it is randomly selected. In addition, the observations are not
     influenced by each other, so they are independent. Finally, from the data above, it meets
      the assumptions of at least 10 successes and 10 failures in each group, so the sample is
      big enough to make inference. Therefore, none of them seems to be violated.

24. D. Here, we tried to compare two independent proportions.

25. B. In this research, we would like to know if male professionals are more likely
      to have children than female professionals of the same age (33-38). So, we
      would only use the data: women aged 33-38 vs. men aged 33-38.

26. D. Since the sample is randomly selected, the observations are independent, the
      Data is categorical, and the sample size is big enough, there's no problem to violate the
      assumptions.

27. C. Use the rule p-value <alpha, reject H_o. Our reasonable alpha levels
       are .10,.05, and .01. We reject H_0 at all these levels. So, we can
    conclude that professional men are significantly more likely to have children
       than professional women in this age group.

28. A. Since the CI includes 0, we cannot conclude more Americans support space
      advances than a solution to world hunger.

29. A. The data is continuous, so it has to be a mean problem. There is only one
      group, so that it is a one sample mean problem. We want to make inference
        for the mean of a population.

30. D. The data is continuous (age), so that it is a mean problem. There are two
      independent groups (same gender and different gender), so it is a two sample
      mean problem.

31. C. The data is continuous (yield), so it has to be a mean problem. In a matched

pairs study, subjects are matched in pairs and the outcomes are compared within each matched pair.

32. B. The data is categorical, (yes/no), so it is a proportion problem. There is one group (the 50 shoppers), so it is a one sample proportion problem.

33. C. The data (distance) is continuous, so it is a mean problem. In a matched pairs study, subjects are matched in pairs and the outcomes are compared within each matched pair.

34. E. The data is categorical (yes/no), so it is a proportion problem. Here, political action group want to compare two groups (with and without college degrees). So, it is a two sample proportion problem.

35. A. The data is continuous (amount of ice cream), so it is a mean problem. There is also only one sample of 20. One sample mean.

36. C. The data is continuous (productivity), so it is a mean problem. Keyword: "same".

37. E. The data is categorical (male/female), so it is a proportion problem. There are two years (1982 & 2001), so this is a two independent proportion problem.

38. E. Since the P-value is a probability, so it must be between 0 and 1.

39. D. A confidence interval gives a region where the unknown parameter lies. So, when the interval includes the hypothesized value of the parameter, it is possible for that value to be equal to the parameter. Therefore, we definitely don't have evidence to show that the alternative is true, so fail to reject H_o.

40. B. We can find the answer from the Minitab output. (parkwkda)

41. A. We want to compare the means (the average time) of two independent groups (Saturday and weekdays).

42. C. The p-value for 2 independent means (parkwkday vs. parksat) is .0005. The p-value is so small that we reject the H_o. There is evidence that it takes longer on Saturday to park.

43. A. We want to compare the means (the average time) of two independent groups (Saturday and Sunday). We want to compare the means (the average time) of two independent groups (Sat. w/games) and (Sat. without games). This is not matched pairs because we are not requiring the same people to park on both days.

44. A. The p-value for 2 independent means (parksat & parkgameday) is .23. This p-value is big. Fail to reject H_o. There is not enough evidence to show a difference.

45. A B is false. First of all, population does not have to be normal. Small departures from Normality are o.k. Second of all, keep in mind that we do not know σ most of the time, so we cannot use the standard normal distribution. We use the t distribution instead. C is false because neither of the two methods can be used for categorical data.

46. A A is false, because although matched pairs design uses dependent samples, it is still preferred. For reasons described in choice B and C. Therefore, B and C are true, A is not.

47. A Note that our claim states "are likely to spend less". Therefore, we should go with H_a: $\mu < 942$. Ho should then be Ho: $\mu = 942$.

48. A $$TS = \frac{estimator - \#Ho}{s\,t\,err} = \frac{820.24 - 942}{260.4 / \sqrt{25}} = -2.338$$

We will use std.error instead of the term estimate of standard error for this material, since all standard errors are estimates in this section.

49. A Explanation: We are going to look up 2.338 on the t-table. Since the t-distribution is symmetric, $P(t < -2.338) = P(t>2.338)$, which is exactly the upper tail probability given by the table. Looking at D.F. 24 (n - 1 = 25 - 1 = 24), we see that 2.064 corresponds to P = 0.025, whereas 2.492 corresponds to P = 0.01. Since 2.064 < 2.338 < 2.492, we conclude that 0.01 < P value < 0.025. Since this is a one sided test, we do not need to multiply this range by 2.

50. C $Ho: \mu = 50$ vs. $Ha: \mu < 50$
$$TS = \frac{estimator - \#Ho}{s\,t\,err} = \frac{48.4 - 50}{7.5 / \sqrt{100}} = -2.13$$
Because n is greater than 30, we can use the z table. The probability less than -2.13 is 0.0166.

51. B
$$TS = \frac{estimator - \#Ho}{s\,t\,err} = \frac{1.4 - 1.3}{0.3 / \sqrt{25}} = 1.6667$$
The degree of freedom associated with this test are n-1 = 24
From the t-table:
P (T-stat $\neq$ 1.67) = 2*P (T-stat > 1.67)   (Two-sided test)
P (T-stat > 1.67) is between .05 and .1
Therefore, P (T-stat $\neq$ 1.67) is between .1 and .2

52. A Looking at the t-table we find that the upper tail probability matches exactly a p-value of .025,
but this is two sided test therefore the p-value is 2*.025 = .05
P (T $\neq$ 2.11) = 2*P (T>2.11) = 2*.025 = .05

53. C   Degrees of freedom = n-1 = 9

54. B   The definition of the p-value is the probability that you would see a test statistic this extreme if the null hypothesis were true.

55. B   The test statistic is of the form $\frac{estimator - \# Ho}{stderr}$. For this problem, the estimator is x-bar, which is not given. The #in Ho is 75 and the stderr is $\frac{s}{\sqrt{n}} = \frac{9}{\sqrt{25}}$.

56. D   Rule: p-value < alpha Reject Ho Since the p-value is between 0.005 and 0.001, we would reject Ho at all values of α (0.10, 0.05 and 0.01).

57. C   Look at the output in between the parenthesis. It says (VS not =). This means that the alternative hypothesis is "not equal to".

58. D   The men and women are from two independent samples and we are testing the average amount spent on books so this is a comparison of two independent means.

59. D   Rule: p-value < alpha Reject Ho Since the p-value is 0.076, we would reject Ho at only α equal to 0.10 not 0.05 or 0.01.

60. A   Since the alternative is $\mu_D > 0$, we are trying to show that the difference is positive, implying that the average height after sleep is greater than the average height before sleep.

61. C   Since the p-value is greater than 0.05, we cannot reject the null hypothesis. This implies that there is no evidence to make a conclusion that the difference in heights is positive.

62. D   The claim that the researcher wants to test becomes the alternative Ha: μ_D (= weight after sleep – weight before sleep) < 0; for this alternative, the null is Ho: μ_D = 0, implying that the article stated that there is no difference between people's average weight before and after sleep.

63. B   The p-value for the test is less than α, supporting the alternative hypothesis, meaning that the researcher has managed to show that the students at her university weigh less after a night's sleep.

64. C   Both tests were comparisons of two dependent means because in both cases two measurements were performed on the same experimental unit.

65. B   66. A   67. D   68. F   69. C   70. E

71. B   We are interested in determining if it is different from 5. 5.39 is the sample mean.

72. C   We use the t distribution because the sample size is less than 30. The formula for the t-
statistic is $t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n}}} = \dfrac{5.39 - 5}{1.767/\sqrt{25}} = 2.54$

73. B   df=n-1=24 , t= 2.542 .
The probability above 2.542 is between 0.01 and 0.005. So, the probability above -
2.542 is also between 0.01 and 0.005. So, the p-value is those areas added to 0.02 and
0.01

74. A   The formula for the confidence interval is $\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \dfrac{s}{\sqrt{n}}$

Calculation : 5.39- (0.767/√25)*2.064=5.07
5.39+ (0.767/√25)*2.064=5.71

75. C   We have $p < \alpha$ we consequently reject the null hypothesis.

76. C   We are only computed two sided test for McNemar's test

77. C   This is by the definition of McNemar test.
78. A   Dependent since the sample is the same 1600 people where asked both questions.

79. C   Test Statistic= $\dfrac{150 - 86}{\sqrt{150 + 86}} = 4.17$ or $\dfrac{86 - 150}{\sqrt{150 + 86}} = -4.17$

80. D   Since the p-value for the test is very small, p-value <0.0001, you would reject Ho.

81. A   By definition, a Type I error results from Rejecting Ho when it is true.

82. D   By definition, a Type II error results when failing to Reject Ho when it is Ho false.